

Matrix Theory for Minimal Trellises

Iwan M. Duursma *

September 28, 2015

Abstract

Trellises provide a graphical representation for the row space of a matrix. The product construction of Kschischang and Sorokine builds minimal conventional trellises from matrices in minimal span form. Koetter and Vardy showed that minimal tail-biting trellises can be obtained by applying the product construction to submatrices of a characteristic matrix.

We introduce the unique reduced minimal span form of a matrix and we obtain an expression for the unique reduced characteristic matrix. Among new properties of characteristic matrices we prove that characteristic matrices are in duality if and only if they have orthogonal column spaces, and that the transpose of a characteristic matrix is again a characteristic matrix if and only if the characteristic matrix is reduced. These properties have clear interpretations for the unwrapped unit memory convolutional code of a tail-biting trellis, they explain the duality for the class of Koetter and Vardy trellises, and they give a natural relation between the characteristic matrix based Koetter-Vardy construction and the displacement matrix based Nori-Shankar construction.

For a pair of reduced characteristic matrices in duality, one is lexicographically first in a forward direction and the other is lexicographically first in the reverse direction. This confirms a conjecture by Koetter and Vardy after taking into account the different directions for the lexicographical ordering.

1 Introduction

A trellis is a directed graph with edge-labeled paths that represent the possible output sequences of an encoder. Trellises facilitate the decoding, via shortest path algorithms, of sequences that have been received over a noisy communication channel. Important examples of such algorithms are the Viterbi maximum likelihood decoder and the BCJR maximum a posteriori decoder. For linear block codes, encoding is given by matrix multiplication and a trellis represents the vectors in the row space of the matrix.

*University of Illinois at Urbana-Champaign, Dept of Mathematics and Coordinated Science Laboratory. Partially supported by a grant from the Simons Foundation (#280107). duursma@math.uiuc.edu.

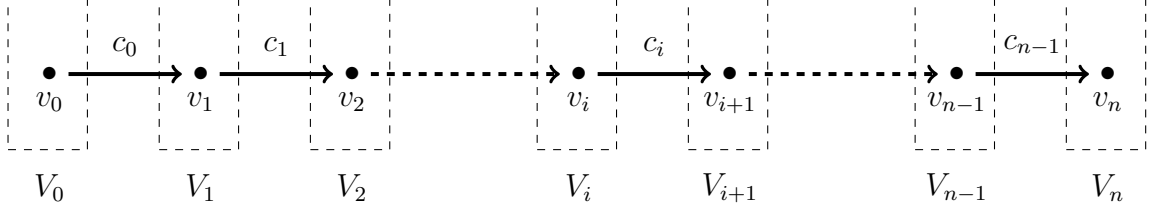


Figure 1: Trellis path for a vertex partition $V_0 \cup V_1 \cup \dots \cup V_n$

We assume throughout that trellises have no multiple edges and in particular that a path in the trellis is uniquely determined by specifying the vertices along the path. A conventional trellis comes with a partition $V_0 \cup V_1 \cup \dots \cup V_n$ of the vertices, such that both V_0 and V_n are singletons. In the more general setting of tail-biting trellises, V_0 and V_n are in bijection. A path (v_0, v_1, \dots, v_n) is tail-biting if the bijection between V_0 and V_n matches the head v_n of the path to its tail v_0 . It is common use to identify V_0 and V_n and to refer to a tail-biting path as a cycle with origin v_0 or cycle for short. A tail-biting path with edge labels c_0, c_1, \dots, c_{n-1} is said to represent the row vector $(c_0, c_1, \dots, c_{n-1})$.

Let $T \subset V_0 \times V_1 \times \dots \times V_n$ be the full collection of tail-biting paths in a trellis. A trellis is said to represent the row space C of a matrix if each row vector represented by a path in T belongs to C and if each row vector in C is represented by a path in T . The representation is one-to-one if each row vector in C is represented by a unique path in T . The trellis construction problem is to find minimal trellises that represent the row space of a given matrix one-to-one while minimizing the vertex count or any other relevant complexity measure for the trellis.

Constructions for linear trellises

All trellises considered in this paper are linear trellises. For a linear trellis that represents the row space of a matrix with coefficients in a field F , the vertex subsets V_i are F -linear spaces. A tail-biting path has a vertex labeling $\nu = (v_0, v_1, \dots, v_n) \in T$ and an edge labeling $c = (c_0, c_1, \dots, c_{n-1}) \in C$. The vertex-edge labelings (ν, c) form a linear subspace $S(T) \subset T \times C$ called the label code of the trellis. A linear trellis is determined by specifying a basis for the label code, that is by specifying the vertex-edge labelings for a subset of paths that generates the trellis as F -linear space. Two trellises are linearly isomorphic if there exist linear isomorphisms $\alpha_i : V_i \rightarrow V'_i$, for $0 \leq i \leq n$, that are compatible with edges labelings, i.e. that map edges (v_i, c_i, v_{i+1}) to edges $(\alpha_i(v_i), c_i, \alpha_{i+1}(v_{i+1}))$. Figure 2 illustrates for two different trellis constructions how they lead to different but isomorphic trellises (over the ternary field $F = \{0, 1, 2\}$). We briefly describe each of the two constructions.

The product construction uses a basis of step functions for the vertex labeling. For a path (ν, c) in the basis of the label code every vertex along the path is labeled either 0 or e_ν , such that the set of labels $\{e_\nu\}$ for the paths in the basis is linearly independent. There are

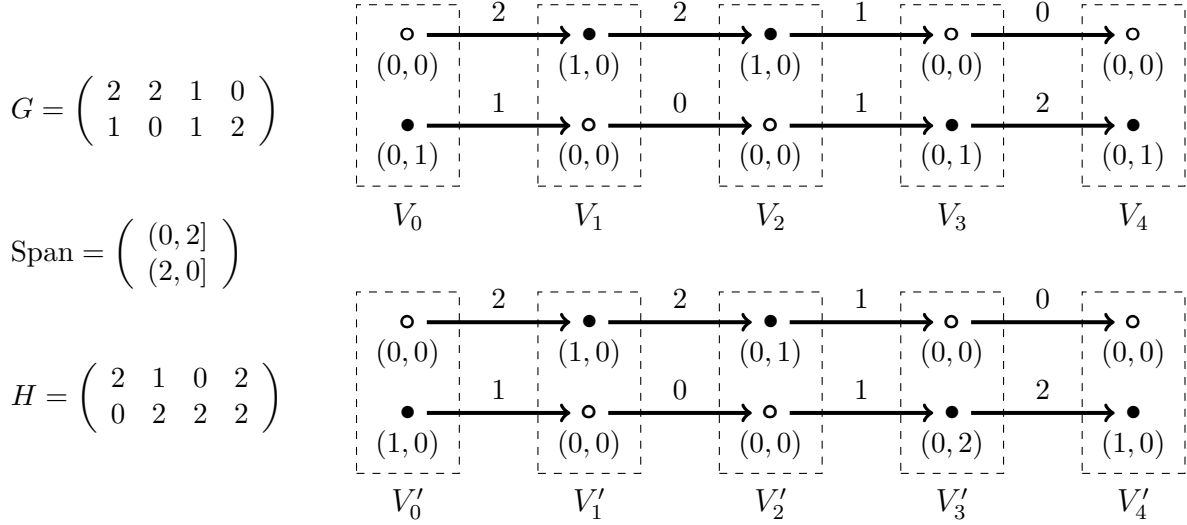


Figure 2: Two isomorphic trellis representations for the row space of the matrix G . Product construction (above) and BCJR construction (below)

precisely two changes along the path: $v_i = 0$ and $v_{i+1} = e_\nu$ for a unique i , and $v_j = e_\nu$ and $v_{j+1} = 0$ for a unique j (an allzero path will not be part of the basis and the special case of a never zero path is needed only for disconnected trellises). The transitions are determined by choosing a span $(i, j]$ for the path such that $c_i, c_j \neq 0$ and $c_k = 0$ whenever k is *not in the span*.¹ The span of a path is in general not unique. A Koetter-Vardy trellis is the special case where the selected generators have a span $(i, j]$ that is minimal (among all spans of the form $(i', j]$ or $(i, j']$). In general a trellis has precisely n such minimal spans $(i, j]$, each with a unique $i \in \{0, 1, \dots, n-1\}$ and a unique $j \in \{0, 1, \dots, n-1\}$. The product construction was introduced for conventional trellises in [15].

The vertex labeling in the tail-biting version of the BCJR construction depends only on the labeling of the vertices in V_0 . The labeling of the other vertices along a path follows the rule $v_{i+1} = v_i + c_i h_i$ of the classical BCJR construction, for parity-check vectors h_i . Note that $v_n = v_0 + \sum c_i h_i$. To guarantee that the trellis represents the linear space C , the h_i are chosen such that $\sum c_i h_i = 0$ if and only if $c \in C$. A span based BCJR trellis is the special case where the label v_0 in a path is decided by the rule $c_k = 0$ whenever k is *not in the span* together with the rule $v_{i+1} = v_i + c_i h_i$. The tail-biting versions of the BCJR construction were introduced by Nori and Shankar. Their formulation emphasizes coset decompositions of the row space. Coset decompositions were introduced for conventional trellises in [12].

The two constructions give two ways to describe minimal trellises.

$$\{\text{minimal trellises}\} \subset \{\text{KV-trellises}\} \subset \{\text{span based BCJR trellises}\}$$

¹ $k \notin (i, j]$ if the ordered triple i, j, k is even, i.e., $i < j < k$, $k < i < j$ or $j < k < i$.

The inclusions hold for trellises up to isomorphism. Thus a KV-trellis can be constructed as a span based BCJR trellis up to isomorphism, as in Figure 2. The first inclusion is proved in [14], [10], and the second in [17], [10]. The definitions in [10], which we follow for this paper, differ slightly from those in [14] and [17].

Koetter and Vardy [14] define a characteristic matrix X for a given matrix G as a square matrix with the same row space as G such that the row spans start and end in different positions. The matrices G and H in Figure 2 have characteristic matrices X and Y (the rows are ordered such that spans of rows in X start on the diagonal and spans of rows in Y end on the diagonal).

$$X = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 2 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (1)$$

In general, the characteristic matrix is not unique (even after ordering and normalizing). For any minimal trellis, there exists a characteristic matrix such that the trellis can be constructed as a product trellis from rows in the characteristic matrix ([14], [10]). The product trellis in Figure 2 is a minimal trellis and was constructed using rows 1 and 3, and their spans $(0, 2]$ and $(2, 0]$, in the matrix X . While characteristic matrices are not unique, there exists a bijection between the characteristic matrices for a pair of matrices G and H with maximal orthogonal row spaces such that trellises constructed from matching characteristic matrices X and Y are in duality [11].

Main results

The main results are presented in Sections 4 and 5. There we define a characteristic matrix in a novel way and from the definition we easily obtain its main properties including a straightforward description of the relation between a characteristic matrix and its dual. An important part of our definition is the notion of a reduced minimal span form, which is developed in Sections 2 and 3.

Minimal span matrices (the subject of Sections 2 and 3) are characterized by the left-right property: no two rows start or end in the same positions. In other words, only row exchanges are needed to bring the matrix in either left or right echelon form. Their connection to trellises is two-fold. The product construction applied to the rows of a matrix yield a minimal trellis if and only if the matrix is in minimal span form. Secondly, the vertex count of a minimal conventional trellis can be read off easily from the shape of a matrix in minimal span form. Throughout the literature, in many examples and trellis descriptions, it is noted that the minimal span form is in general not unique and it appears that some arbitrary choices need to be made to settle for a particular minimal span form. In this paper we bring a matrix in reduced minimal span form in the same way that back substitution brings a matrix from echelon form into reduced echelon form. This reduction is equivalent

to reducing an arbitrary LPU decomposition of a nonsingular matrix to a unique reduced LPU decomposition. In this form our proposed reduction is the same as the one used for reduced Bruhat decompositions in [7].

In Section 4 we adopt a definition of characteristic matrices that refers to a matrix in minimal span form. The notion of reduced minimal span form thus allows us to define a unique reduced characteristic matrix. We use two different forms for a characteristic matrix, a left-ordered form (controller form) and a right-ordered form (observer form). Let G and H be the pair of orthogonal matrices from Figure 2, with characteristic matrices given by (1). For the controller form of the characteristic matrix X for G , we write X as the sum of an upper triangular matrix and a lower triangular matrix.

$$X = X_0 + X_1 = \begin{bmatrix} 2 & 2 & 1 & 0 \\ \cdot & 1 & 1 & 1 \\ \cdot & \cdot & 1 & 2 \\ \cdot & \cdot & \cdot & 2 \end{bmatrix} + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ 1 & 0 & \cdot & \cdot \\ 2 & 1 & 0 & \cdot \end{bmatrix}$$

The matrices G and X share the same row space. Since $XH^T = 0$, $(X_0|X_1)(H|H)^T = 0$. In general we derive X_0 and X_1 from a reduced minimal span form for the null space of $(H|H)$. The dual characteristic matrix Y is decomposed as follows.

$$Y = Y_1 + Y_0 = \begin{bmatrix} \cdot & 0 & 1 & 2 \\ \cdot & \cdot & 0 & 1 \\ \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} + \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 1 & 2 & \cdot & \cdot \\ 1 & 1 & 2 & \cdot \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

It satisfies $YG^T = 0$ and thus $(Y_1|Y_0)(G|G)^T = 0$. Similar to the above, we derive Y_1 and Y_0 using a reduced minimal span form for the null space of $(G|G)$. In Section 5, we express duality for characteristic matrices as

$$\begin{pmatrix} X_0 & X_1 \\ 0 & X_0 \end{pmatrix} \begin{pmatrix} Y_0 & 0 \\ Y_1 & Y_0 \end{pmatrix}^T = \begin{pmatrix} -I & I \\ 0 & -I \end{pmatrix}.$$

In Theorem 5.5 we prove, avoiding some of the technical details of earlier proofs, that KV-trellises defined with X and with Y are in duality. Furthermore, in Theorem 5.10 we show that two characteristic matrices are in duality if and only if they have orthogonal column spaces. And in Theorem 5.17 that the transpose of a characteristic matrix is again a characteristic matrix if and only if the characteristic matrix is reduced.

Originally Koetter and Vardy proposed that a duality between characteristic matrices holds for the lexicographically first characteristic matrices. In [10] examples were found where this was not the case. In Theorem 5.8 we adjust the original conjecture and prove that reduced characteristic matrices are in duality. Reduced characteristic matrices are lexicographically first when read out from right to left (for X) or left to right (for Y). After taking into account the different directions of the time axis in a trellis and its dual the original conjecture appears correct.

To describe the KV-trellises both as a product trellis and as a BCJR trellis, we use the column space of Y as labels for trellises constructed with rows in X .

$$Y^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

This leads to label codes of the form

$$S(Y^T|X) = \left[\begin{array}{c|c|c|c|c|c|c|c} 0 & 0 & 0 & 0 & 2 & 1 & 1 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 1 & 2 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 1 & 2 & 2 & 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 2 & 0 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 0 & 1 & 0 \end{array} \right]$$

We prove in Section 5.3 that the following properties hold in general. The single columns in the matrix lie in the null space of the square matrices (Lemma 5.10). Any two consecutive square matrices differ by a rank-one matrix (Lemma 5.12). And each of the square matrices contains a basis in echelon form for the row space Y^T (Theorem 5.13). Moreover, by Theorem 5.5 the square matrices in the dual label code $S(X^T|Y)$ relate to those of $S(Y^T|X)$ by transposition (up to a sign change and only off the diagonal).

2 Minimal span form of a matrix

The span (or support interval) of a nonzero vector $c = (c_0, c_1, \dots, c_{n-1})$ is the largest interval $[i_0, i_1] \subset [0, n-1]$ with $c_{i_0} c_{i_1} \neq 0$. We say the vector starts in i_0 and ends in i_1 . The size $i_1 - i_0$ of the span is called the spanlength of the vector. The spanlength of a matrix is the sum of the spanlengths of its rows. A matrix is in minimal span form if its spanlength is minimal among all row equivalent matrices. An easy way to decide if a matrix is in minimal span form is via the *left-right property*.

Lemma 2.1 ([15], [16]). *The following are equivalent.*

- (1) *A matrix is in minimal span form.*
- (2) *No two distinct rows in the matrix start in the same position or end in the same position.*

Proof. The proof is short and elementary. If (2) fails for two rows in a matrix then we can reduce the spanlength of one of the two rows and (1) fails. If (2) holds for a matrix then the rows of the matrix are in row echelon form after reordering and the starting positions agree with the column indices for the left pivots. By the same argument the ending positions agree with the column indices for the right pivots. Thus the spanlength is the same for all matrices with (2). By the first part of the proof this spanlength is that of a matrix in minimal span form. \square

Let I_0 be the collection of left pivot positions in a matrix and let I_1 be the collection of right pivots positions. From the proof of the lemma we see that the spanlength for a matrix in minimal span form is given by $\sum_{I_1} i_1 - \sum_{I_0} i_0$.

Lemma 2.2 ([15], [16]). *Any two minimal span forms for a matrix have the same collection of spans.*

Proof. The proof of the previous lemma shows that two minimal span forms share the same set I_0 of starting indices. For a given starting index $i_0 \in I_0$ both minimal span forms contain a row that starts in i_0 . Let the two rows have spans $[i_0, i_1]$ and $[i_0, i'_1]$ and assume that $i_1 < i'_1$. Then the row with span $[i_0, i'_1]$ can be reduced to a row with smaller span $[i'_0, i'_1]$. The reduced row is still independent of the other rows since i'_1 has not changed and this contradicts that the matrix was in minimal span form. Thus $i_1 = i'_1$ as required. \square

The left-right property in Lemma 2.1 expresses that after ordering the rows of a matrix in minimal span form we may assume that the minor in the positions I_0 (the leading nonsingular full minor in the matrix) is upper triangular and that the minor in the positions I_1 (the trailing nonsingular full minor in the matrix) is lower triangular up to row permutations. Denoting these two minors by U and PL , respectively, the spans are determined by the permutation matrix P . Lemma 2.2 expresses that $P = P'$ for any two row equivalent matrices $[U|PL] \sim [U'|P'L']$. We relate this last property to the LPU -decomposition of an invertible matrix A .

Lemma 2.3. (*LPU-decomposition*) *Let A be an invertible matrix. Then there exists a decomposition $A = LPU$ with L lower triangular, P a permutation matrix and U upper triangular. The decomposition is in general not unique but the matrix P is uniquely determined by the matrix A .*

Proof. The result is well known.² To see how the uniqueness claim

$$LPU = L'P'U' \Rightarrow P = P' \quad (2)$$

follows from the previous lemmas, note that

$$LPU = L'P'U' \Leftrightarrow [U|P^{-1}L^{-1}] \sim [U'|P'^{-1}L'^{-1}]$$

and apply Lemma 2.2,

$$[U|P^{-1}L^{-1}] \sim [U'|P'^{-1}L'^{-1}] \Rightarrow P = P'. \quad (3)$$

\square

²But should not be confused with the different decomposition $PA = LU$.

2.1 Bruhat decomposition

The uniqueness claim (2) is an instance of a modified Bruhat decomposition for the general linear group. In general, a Bruhat decomposition of a (connected, reductive) linear algebraic group is of the form $G = \cup_{w \in W} BwB$, for a Borel subgroup B and a Weyl group W . The double coset decomposition $\cup_{w \in W} BwB$ partitions the group G and associates to each group element g a unique element $w \in W$ (e.g. [20, Theorem 10.2.7 (“Bruhat’s lemma”)]). For the general linear group, the Bruhat decomposition uses for B the group of invertible upper triangular matrices and for W the group of permutation matrices.

The standard Bruhat decomposition assigns to a matrix A a unique permutation matrix $U_1^{-1}AU_2^{-1} = P$. In general it is possible to reduce an invertible matrix A to a unique permutation matrix P using row operations and column operations of any fixed triangular type. The outcome P depends on the chosen types but not on the sequence of reduction operations. To illustrate that the outcome depends on the chosen types we reduce the same binary matrix A in two different ways to a permutation matrix, via a standard Bruhat decomposition and via a LPU -decomposition.

$$\begin{aligned} U_1^{-1}A &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} U_2. \\ L^{-1}A &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} U. \end{aligned}$$

To interpret the matrix P assigned to A by the LPU decomposition, define a leading submatrix of a matrix as a submatrix positioned in the north-west corner of the matrix. The following is clear and stated without proof.

Lemma 2.4. *Let A be any matrix. For any lower triangular invertible matrix L and any upper triangular invertible matrix U , leading submatrices in A and LAU of the same size are of the same rank.*

It follows that the matrix $P = L^{-1}AU^{-1}$, that is assigned to A by an LPU -decomposition, describes the rank of any leading submatrix in A by counting the number of nonzero entries in the corresponding leading submatrix of P . In particular the matrix P is unique, which confirms in a different way the claim (2). Similarly, for each of the four corners of the matrix, we can reduce A to a permutation matrix P that gives the rank of submatrices positioned in that corner:

$$\begin{aligned} P_{NW} &= L^{-1}AU^{-1}, & P_{NE} &= L^{-1}AL^{-1} \\ P_{SW} &= U^{-1}AU^{-1}, & P_{SE} &= U^{-1}AL^{-1}. \end{aligned}$$

Lemma 2.5. *Let $PT = T'P'$, for invertible triangular matrices T and T' and for permutation matrices P and P' . Then $P = P'$.*

Proof. Submatrices of P and $P' = T'^{-1}PT$ of the same size in one of the four corners (which corner depends on the type of T and T' , upper or lower triangular) are of the same rank and thus $P = P'$. \square

3 Reduced minimal span form of a matrix

Let G be a given matrix. As in the previous section, let I_0 be the collection of left pivot positions and let I_1 be the collection of right pivot positions. As pointed out after Lemma 2.2 the rows in a minimal span form can be ordered such that the matrix is in echelon form on the left.

$$G|_{I_0} = U \quad \text{and} \quad G|_{I_1} = PL. \quad (4)$$

We will call a minimal span form of this type left-ordered. An example is the binary matrix

$$G = \left(\begin{array}{ccc|cc} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right) = \left(\begin{array}{cc|ccc} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right)$$

Among all minimal span forms (4) we select a unique reduced minimal span form.

Lemma 3.1. *There exists a minimal span form of the form (4) with the additional property that the trailing nonzero element in each row is the first nonzero element in its column.*

Proof. Let $(i, \pi(i))$, $1 \leq i \leq n$, be the nonzero entries in P , for a permutation π of $\{1, 2, \dots, n\}$. Thus $\pi(i)$ marks the last column in row i of PL with a nonzero entry. The nonzero entries in column $\pi(i)$ with row index $i' < i$ have $\pi(i') > \pi(i)$. The pair i, i' corresponds to an inversion of π . Using a sequence of elementary row operations in upper triangular form we can clear all nonzero entries in column i with row index $i' < i$. To assure that columns that have been cleared are not affected by clearing other columns, columns should be cleared in reverse order, from last to first column. \square

A matrix PL has the property in the lemma if and only if $PL = L'P'$ for a lower triangular matrix L' and a permutation matrix P' . By Lemma 2.5 we may assume that $P' = P$.

Definition 3.2. A minimal span form is called left-ordered and right-reduced if

$$G|_{I_0} = U \quad \text{and} \quad G|_{I_1} = PL = L'P. \quad (5)$$

It is upper triangular in the leading pivot columns and lower triangular up to a row permutation in the trailing pivot columns. Moreover it is lower triangular up to a column permutation in the trailing pivots. Less formal, a minimal span form is reduced if it is upper triangular on the leading pivots (and thus zero in the positions west and south of the leading pivots) and zero in the positions east and north of the trailing pivots.

Proposition 3.3. *Every matrix has a unique reduced minimal span form.*

Proof. Lemma 3.1 proves the existence of a minimal span form G that satisfies (5). Lemma 2.2 shows that the matrix P is the same for any two minimal span forms. Thus, the matrix $U^{-1}L'$ is the same for any two reduced minimal span forms. It follows that U and L' are unique (up to a constant diagonal matrix). \square

3.1 Dual minimal span forms

Definition 3.2 characterizes the reduced minimal span form of a matrix via the properties left-ordered and right-reduced. This follows the convention for a reduced row echelon form where the echelon form appears on the left. In both cases forward elimination brings the matrix in echelon form on the left. The back substitution phase of Gaussian elimination clears columns on the left to obtain the reduced row echelon form. For the reduced minimal span form the second phase of clearing columns is carried out on the opposite right side of the matrix.

As in [16, Section IV], the standard row reduced echelon form might be called a “left” reduced row echelon form. If pivot columns are selected from the right, one obtains what might be called a “right” reduced row echelon form. We denote a left reduced echelon form for G by G_0 and a right reduced echelon form by G_1 .

$$G_0|_{I_0} = I, \quad G_1|_{I_1} = I.$$

For a pair of matrices G and H with maximal orthogonal row spaces, the matrix H has leading pivots $J_0 = \{0, 1, \dots, n-1\} \setminus I_1$ and trailing pivots $J_1 = \{0, 1, \dots, n-1\} \setminus I_0$. The echelon forms G_0 and H_1 are in duality via

$$G_0|_{I_0} = I, \quad G_0|_{J_1} = A, \quad H_1|_{I_0} = -A^T, \quad H_1|_{J_1} = I.$$

To obtain a duality for reduced minimal span forms we add to Definition 3.2 a matching reduced minimal span form for the matrix H .

Definition 3.4. A matrix G is in left-ordered right-reduced minimal span form if

$$G|_{I_0} = U \quad \text{and} \quad G|_{I_1} = PL = L'P.$$

A matrix H is in right-ordered left-reduced minimal span form if

$$H|_{I_0} = PU = U'P \quad \text{and} \quad H|_{J_1} = L.$$

The minimal span form for G is normalized by choosing L with 1s on the diagonal. The minimal span form for H is normalized by choosing U with 1s on the diagonal.

We write G_{01} for a reduced minimal span form of the first type and H_{10} for a reduced minimal span form of the second type. We will rarely use the full long form terminology of the definition. Instead we write $G \perp H$ for a pair of matrices with maximal orthogonal row spaces and call the pair reduced if $G = G_{01}$ and $H = H_{10}$. This is illustrated for the pair $G_{01} \perp H_{10}$ below.

$$\begin{aligned}
G_0 &= \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right) & G_{01} &= \left(\begin{array}{ccc|cc} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right) = \left(\begin{array}{cc|ccc} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right) \\
H_1 &= \left(\begin{array}{ccc|cc} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array} \right) & H_{10} &= \left(\begin{array}{ccc|cc} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{array} \right) = \left(\begin{array}{cc|ccc} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{array} \right)
\end{aligned}$$

The property that characterizes a matrix in minimal span form is the LR-property (property (2) in Lemma 2.1). Similarly, we may call the characterizing property of a reduced minimal span form the SW-NE property (after ordering the rows on the left or on the right, entries south and west of a leading pivot are zero and entries north and east of a trailing pivot are zero). This property will have a clear interpretation in the duality of characteristic matrices in Section 4.

3.2 Reduced Bruhat decompositions

Given the connection between minimal span forms and Bruhat decompositions that was pointed out in Section 2.1, we want to interpret the reduced minimal span form of a matrix in terms of Bruhat decompositions. This leads us to the notion of a reduced Bruhat decomposition defined by Kolotriline and Yeremin in [7]. About a Bruhat decomposition $A = V\pi_A U$ (A nonsingular, U and V upper triangular) they write [loc.sit. page 422]: “The triangular factors V and U ... are not in general uniquely determined. To remove this uncertainty, one can consider the so-called reduced Bruhat decomposition.”

Definition 3.5 ([7]). The decomposition $A = V\pi_A U$ will be referred to as the reduced-on-the-right or on-the-left Bruhat decomposition if

$$(\pi_A)U(\pi_A)^T \quad \text{or} \quad (\pi_A)U(\pi_A)^T,$$

respectively, is lower triangular.

Further on [loc.sit. page 423]: “It is easy to show that all factors of the reduced Bruhat decomposition are uniquely determined.” The definition is motivated by applications in numerical analysis [loc.sit page 421]: “the Bruhat decomposition ... allows the formulation of a global analytical criterion for optimizing sparsity patterns of triangular factors ...”. Clearly the reductions applied in Definition 3.4 agree with those of Definition 3.5 when the latter are rewritten for a modified Bruhat decomposition.

The nonuniqueness of matrices L and U in the decomposition $A = LPU$ is due to inversions in the permutation matrix P . This is illustrated by

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} = LPU.$$

To the one-parameter family of LPU-decompositions corresponds the one-parameter family of minimal span forms

$$(A|I) = \left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cc|cc} 1 & a & a & 1 \\ 0 & 1 & 1 & 0 \end{array} \right) = (U|P^{-1}L^{-1})$$

Clearly, the choices for L and U that minimize the number of operations in the assignment $A \mapsto P = L^{-1}AU^{-1}$ (in other words that optimize the sparsity of L^{-1} and U^{-1}) are the choices $L = U = I$ and $a = 0$. Which is our choice for the reduced minimal span form of the matrix $(A|I)$.

4 Characteristic matrix

Koetter and Vardy [14] define a redundant set of generators, called characteristic generators, for the row space of a matrix. From the redundant set it is straightforward to recover not only a minimal span form for the original matrix but also for the cyclic permutations of the original matrix. In [14], the set of characteristic generators is built using an efficient procedure. The procedure checks the minimal span forms for cyclicly permuted versions of the original matrix one cyclic shift at a time and each time decides if a new generator should be added to the list. The final list is shown to be of size n and has the property that the spans for the generators start and end in distinct positions. The main result of [14] is that any reduced linear tail-biting trellis can be built with the product construction from a subset of characteristic generators. For this result, the existence of characteristic generators and the fact that they can be computed efficiently when needed is sufficient. In this section we describe characteristic generators in a different way and establish new properties. The properties lead to short proofs for duality of Koetter-Vardy tail-biting trellises.

Definition 4.1. A row vector $(c_0, c_1, \dots, c_{n-1})$ has span $(i, j]$ if $c_i, c_j \neq 0$ and if $c_k = 0$ whenever k is *not in the span*. Here $k \notin (i, j]$ if and only if the ordered triple i, j, k is even, i.e., $i < j < k$, $k < i < j$ or $j < k < i$. The span of a row vector is in general not unique.

The definition generalizes the definition of conventional span (or support interval) used in Section 2. The length of a conventional span $(i, j]$ is $j - i$ and it is $n - (j - i)$ for its complement $(j, i]$.

Definition 4.2. Let G be a matrix, $C = \text{row } G$ its row space, and n the number of columns. A *set of characteristic generators* for C is a subset of n vectors such that the spans of the vectors start and end in distinct positions and such that the sum of the spanlengths of the vectors is minimal. A square matrix is called a *characteristic matrix* for G if its rows form a set of characteristic generators for the row space of G .

We call X a left-ordered characteristic matrix if rows are ordered such that each row has span starting on the diagonal, and in that case write $X = X_0 + X_1$, for an upper triangular matrix X_0 and a strictly lower triangular matrix X_1 . We call Y a right-ordered

characteristic matrix if rows are ordered such that each row has span ending on the diagonal, and in that case write $Y = Y_1 + Y_0$, for a lower triangular matrix Y_0 and a strictly upper triangular matrix Y_1 . The unwrapped version $(X_0|X_1)$ of the matrix X is in left-ordered minimal span form and the unwrapped version $(Y_1|Y_0)$ of the matrix Y is in right-ordered minimal span form. Circulant matrices are a special case of characteristic matrices. The unwrapping of two circulant matrices is shown below, for polynomials $c(x) = c_0 + c_1x + c_2x^2$ and $d(x) = d_0 + d_1x + d_2x^2$ such that $c(x)d(x) = x^4 - 1$.

$$X = \begin{pmatrix} c_0 & c_1 & c_2 & 0 \\ 0 & c_0 & c_1 & c_2 \\ c_2 & 0 & c_0 & c_1 \\ c_1 & c_2 & 0 & c_0 \end{pmatrix} \quad (X_0|X_1) = \left(\begin{array}{cccc|cccc} c_0 & c_1 & c_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_0 & c_1 & c_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_0 & c_1 & c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_0 & c_1 & c_2 & 0 & 0 \end{array} \right)$$

$$Y = \begin{pmatrix} d_0 & 0 & d_2 & d_1 \\ d_1 & d_0 & 0 & d_2 \\ d_2 & d_1 & d_0 & 0 \\ 0 & d_2 & d_1 & d_0 \end{pmatrix} \quad (Y_1|Y_0) = \left(\begin{array}{cccc|cccc} 0 & 0 & d_2 & d_1 & d_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_2 & d_1 & d_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_2 & d_1 & d_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & d_2 & d_1 & d_0 \end{array} \right)$$

The spanlength of a vector before unwrapping agrees with the conventional spanlength after unwrapping. Let $G \perp H$ be a pair of matrices with maximal orthogonal row spaces, let X be a left-ordered characteristic matrix for G and Y a right-ordered characteristic matrix for H . Unwrapping of the rows in X yields a matrix $(X_0|X_1) \perp (H|H)$. Since X_0 is invertible (its rows start in distinct positions), the two matrices

$$\left(\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G \end{array} \right) \perp (H|H) \quad (6)$$

have maximal orthogonal row spaces. Similarly, the matrix Y_0 in the unwrapped version $(Y_1|Y_0)$ of Y is invertible and the matrices

$$(G|G) \perp \left(\begin{array}{c|c} H & 0 \\ \hline Y_1 & Y_0 \end{array} \right) \quad (7)$$

have maximal orthogonal row spaces.

Lemma 4.3. *Let $G \perp H$ be a pair of matrices with maximal orthogonal row spaces.*

- (X) *A matrix $X = X_0 + X_1$ is a left-ordered characteristic matrix for G if and only if the rows of $(X_0|X_1)$ together with those of $(0|G)$ form a matrix in left-ordered minimal span form orthogonal to $(H|H)$.*
- (Y) *A matrix $Y = Y_1 + Y_0$ is a right-ordered characteristic matrix for H if and only if the rows of $(H|0)$ together with those of $(Y_1|Y_0)$ form a matrix in right-ordered minimal span form orthogonal to $(G|G)$.*

Proof. Minimizing the span of x is the same as minimizing the conventional span of $(x_0|x_1)$. The minimum for the latter is attained for a matrix in minimal span form. \square

For the special choices of a reduced minimal span form in the lemma, we obtain what we will call a reduced characteristic matrix.

Definition 4.4. Let $G \perp H$ as in the lemma.

(X) The matrix $X = X_0 + X_1$ is a left-ordered, right-reduced characteristic matrix for G if

$$\left(\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G_{01} \end{array} \right) = \left(\begin{array}{c|c} -I & I \\ \hline 0 & G \end{array} \right)_{01}.$$

(Y) The matrix $Y = Y_1 + Y_0$ is a right-ordered, left-reduced characteristic matrix for H if

$$\left(\begin{array}{c|c} H_{10} & 0 \\ \hline Y_1 & Y_0 \end{array} \right) = \left(\begin{array}{c|c} H & 0 \\ \hline I & -I \end{array} \right)_{10}$$

Example 4.5.

$$\begin{aligned} \left[\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G \end{array} \right] &= \left[\begin{array}{cccc|cccc} 2 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 1 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 2 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 2 & 2 & 1 & 0 & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & 2 & 2 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 \end{array} \right] \\ &\perp \left[\begin{array}{cccc|cccc} 1 & 1 & 2 & 0 & 1 & 1 & 2 & \cdot \\ \cdot & 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{array} \right] = [H | H]. \\ \\ [G | G] &= \left[\begin{array}{cccc|cccc} 2 & 2 & 1 & 0 & 2 & 2 & 1 & \cdot \\ \cdot & 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{array} \right] \\ &\perp \left[\begin{array}{cccc|cccc} 1 & 1 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 0 & 1 & 2 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 1 & 1 & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 1 & 1 & 2 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 & 1 & 1 \end{array} \right] = \left[\begin{array}{c|c} H & 0 \\ \hline Y_1 & Y_0 \end{array} \right] \end{aligned}$$

Let G have left pivot columns I_0 and right pivot columns I_1 . So that H has left pivot columns $J_1 = \{0, 1, \dots, n-1\} \setminus I_1$ and right pivot columns $J_0 = \{0, 1, \dots, n-1\} \setminus I_0$.

Proposition 4.6. *Let $G \perp H$ as in Lemma 4.3.*

- (X) *A matrix $X = X_0 + X_1$ is a left-ordered characteristic matrix for G if and only if the rows of $(X_0|X_1)$ are orthogonal to $(H|H)$ and have spans starting in distinct elements of $\{0, \dots, n-1\}$ and ending in distinct elements of $I_1 \cup n + J_1$.*
- (Y) *A matrix $Y = Y_1 + Y_0$ is a right-ordered characteristic matrix if and only if the rows of $(Y_1|Y_0)$ are orthogonal to $(G|G)$ and have spans starting in distinct elements of $I_1 \cup n + J_1$ and ending in distinct elements of $n + \{0, \dots, n-1\}$.*

Proof. The matrix $(H|H)$ has spans starting in J_1 . The orthogonal row space in minimal span form therefore has spans ending in $I_1 \cup n + \{0, \dots, n-1\}$. With rows in $(0|G)$ ending in $n + I_1$, the orthogonal row space is in minimal span form if and only if the rows in $(X_0|X_1)$ end in $I_1 \cup n + J_1$. The proof for (Y) is similar. \square

Corollary 4.7. *It follows that*

- (X) *The total spanlength of $(X_0|X_1)$ and thus of X is $|J_1|n = (\text{rank } H)n$.*
- (Y) *The total spanlength of $(Y_1|Y_0)$ and thus of Y is $|I_1|n = (\text{rank } G)n$.*

Definition 4.4 refers to the minimal span forms of $(H|H)^\perp$ and $(G|G)^\perp$. A characteristic matrix can also be obtained directly from minimal span forms for the matrices G and H . This characteristic matrix will in general not be reduced. The rows of a matrix G in minimal span form provide rows for $(X_0|X_1)$ with spans starting in I_0 and ending in I_1 . The remaining generators, with span starting in J_0 and ending in $n + J_1$, can be obtained from a matrix H in minimal span form using the following lemma. The dual construction yields rows for $(Y_1|Y_0)$ with span starting in I_1 and ending in $n + I_0$.

Lemma 4.8. *Let $G \perp H$ be a pair of matrices with maximal orthogonal row spaces and assume that both matrices are in minimal span form.*

- (X) *Let H'_0 be the unique matrix of the same size as H and with support in the trailing pivot columns J_0 of H such that $H'_0 H^T = -I$ and let H'_1 be the unique matrix of the same size as H and with support in the leading pivot columns J_1 of H such that $H'_1 H^T = I$. The rows of $(H'_0|H'_1)$ are orthogonal to $(H|H)$ and have spans starting in J_0 and ending in $n + J_1$.*
- (Y) *Let G'_0 be the unique matrix of the same size as G and with support in the leading pivot columns I_0 of G such that $G'_0 G^T = -I$ and let G'_1 be the unique matrix of the same size as G and with support in the trailing pivot columns I_1 of G such that $G'_1 G^T = -I$. The rows of $(G'_1|G'_0)$ are orthogonal to $(G|G)$ and have spans starting in I_1 and ending in $n + I_0$.*

Proof. Clear. □

Example 4.9.

$$\begin{aligned} H &= \begin{pmatrix} \boxed{\begin{matrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{matrix}} \end{pmatrix} & H'_0 &= \begin{pmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \boxed{\begin{matrix} 1 & 2 \\ 0 & 2 \end{matrix}} \end{pmatrix} & H'_1 &= \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 2 & 1 \end{matrix}} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{pmatrix} \\ G &= \begin{pmatrix} \boxed{\begin{matrix} 2 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{matrix}} \end{pmatrix} & G'_1 &= \begin{pmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \boxed{\begin{matrix} 1 & 2 \\ 0 & 1 \end{matrix}} \end{pmatrix} & G'_0 &= \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 1 & 2 \end{matrix}} \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{pmatrix} \end{aligned}$$

Complementing the rows of G with those of H' we obtain a full set of characteristic generators for G .

Theorem 4.10. *For matrices $G \perp H$ in minimal span form, let G'_0, G'_1, H'_0, H'_1 be as in Lemma 4.8. So that*

$$G'_0 G^T = -I, \quad G'_1 G^T = I, \quad H'_1 H^T = I, \quad H'_0 H^T = -I.$$

Let

$$\begin{aligned} (X_0|X_1) &= \left[\begin{array}{c|c} G & 0 \\ H'_0 & H'_1 \end{array} \right], & X &= X_0 + X_1 = \left[\begin{array}{c} G \\ H' \end{array} \right]. \\ (Y_1|Y_0) &= \left[\begin{array}{c|c} G'_1 & G'_0 \\ 0 & H \end{array} \right], & Y &= Y_1 + Y_0 = \left[\begin{array}{c} G' \\ H \end{array} \right]. \end{aligned}$$

Then X and Y are characteristic matrices for G and H . Moreover,

$$\begin{aligned} X_0 Y_0^T &= -I, & X_0 Y_1^T + X_1 Y_0^T &= I, & X_1 Y_1^T &= 0. \\ Y_0^T X_0 &= -I, & Y_0^T X_1 + Y_1^T X_0 &= I, & Y_1^T X_1 &= 0. \end{aligned}$$

So that $XY^T = 0$ and $Y^T X = 0$.

Proof. Lemma 4.8 in combination with Proposition 4.6 shows that X and Y are characteristic matrices. The two groups of three equalities hold by straightforward verification and adding the equalities in each group yields $XY^T = 0$ and $Y^T X = 0$. □

In [14], the rows of G are called conventional generators and the remaining generators are called circular generators. The latter generators correspond to conventional generators in a cyclicly shifted version of the matrix G . In Lemma 4.8 we rely on the matrix H to construct the circular generators and make no reference to the cyclicly shifted versions of the matrix G . The circular generators correspond to the rows of a matrix H' .

Proposition 4.11. *Let the n rows of $(X_0|X_1)$ be orthogonal to $(H|H)$ with X_0 upper triangular and nonzero on the diagonal and X_1 strictly lower triangular. The following are equivalent.*

- (a) The rows of $(X_0|X_1)$ end in the distinct columns $I_1 \cup n + J_1$.
- (b) The total spanlength of $(X_0|X_1)$ is $|J_1|n$.
- (c) The number of rows with span in $(0, a - 1]$ or in $(a, a + n - 1]$ is $|I_1|$, for each $a \in \{0, 1, \dots, n - 1\}$.
- (d) The number of rows with $(a - 1, a]$ or $(a + n - 1, a + n]$ in their span is $|J_1|$, for each $a \in \{0, 1, \dots, n - 1\}$.

Proof. The equivalence between (c) and (d) is clear, as are the implications (a) implies (b) and (d) implies (b). For $(X_0|X_1)$ such that (b) holds, the rows of $(X_0|X_1)$ are part of a matrix in minimal span form for $(H|H)^\perp$ and, as is in Proposition 4.6, this implies (a). The rows under (c) give an independent set of relations among columns of H . Thus their number is at most $|I_1|$, for each $a \in \{0, 1, \dots, n - 1\}$. The number of rows counted under (d) is therefore at least $|J_1|$, for each $a \in \{0, 1, \dots, n - 1\}$. Equality holds under (c) and (d) for all $a \in \{0, 1, \dots, n - 1\}$ if and only if (b). \square

The proposition may be used to verify that slightly different definitions for the characteristic matrix in [14], [10] and Definition 4.2 are indeed equivalent. By Proposition 4.6, a characteristic matrix X as in Definition 4.2 is characterized by properties (a) and (b). The definitions [14] and [10] make use of properties (c) and (d), respectively. Property (c) is used in [14] to construct a characteristic matrix from conventional generators of cyclicly permuted versions of G . Property (d) is used in [10] as an axiomatic property, from which further properties can be obtained without reference to a specific construction of the characteristic matrix.

Lemma 4.12. *For a characteristic matrix X and for each $a \in \{0, 1, \dots, n - 1\}$, the $|I_1|$ rows of X with span in $(a, a - 1]$ are linearly independent and they form a basis for the row space of X .*

Proof. As in the proof of the proposition, the number of such rows is $|I_1|$ and after unwrapping it is clear that they specify independent relations among the columns of H . \square

The selection of rows in the lemma has an interpretation as a minimal span basis for the matrix G cyclicly shifted by a columns.

Example 4.13. For the pair of binary orthogonal matrices G and H in Section 3, Appendix A gives the reduced characteristic matrices X and Y . The matrices G_{01} and H_{10} appear as submatrices in X and Y . Together they explain all the entries except those in the dashed boxes. The entries in the dashed boxes are in duality with the corresponding full boxes via Lemma 4.8.

4.1 Subspaces defined by intervals

We mention another combinatorial characterization of matrices in minimal span form and characteristic matrices. In each case the rows of the matrix can be used to distinguish between subspaces defined with different support intervals. The situation is as in Boolean Information Retrieval (BIR), where documents can be distinguished based on different matches with a set of index terms. A set of index terms separates the documents in a collection if each document has a unique subset of matching index terms. Similarly, we say that a subset B of vectors separates a collection of subspaces if for any two subspaces V and V' in the family

$$V = V' \Leftrightarrow V \cap B = V' \cap B. \quad (8)$$

We apply this to a matrix G and its row space $C = \text{row } G$. For a subset $I \subset \{0, 1, \dots, n-1\}$ of column indices, let $C(I) = \{c \in C : c_k = 0 \text{ for } k \notin I\}$. We say that the subspace $C(I)$ is supported on an interval if $I = \{i : a \leq i \leq b\}$ for some $0 \leq a < b \leq n-1$.

Lemma 4.14. *The rows in a matrix G in minimal span form separate the collection of subspaces $\{C(I) : I \text{ an interval}\}$.*

Lemma 4.15. *The rows in a characteristic matrix X for G separate the collection of subspaces $\{C(I) : I \text{ an interval}\} \cup \{C(J) : J \text{ a complement of an interval}\}$.*

In the case of G , the rows moreover form a basis for the row space and it follows that the rows have the *subsystem basis property* [9, Theorem 3]. The separating properties of characteristic generators are the same as those of *discrepancies* that are used in [6], [5] to distinguish algebraic functions with a specified behavior at the origin and at infinity. After writing a vector $(c_0, c_1, \dots, c_{n-1})$ as a polynomial $c_0 + c_1x + \dots + c_{n-1}x^{n-1}$, the vector has span $(i, j]$ if and only if the polynomial has a pole of order j at infinity and a zero of order i at the origin. Whereas the characteristic spans of a row space provide information about its trellis structure, the discrepancies of a function field give lower bounds for the minimum distance of codes constructed with the function field [4]. We refer to the listed references for details.

Definition 4.16. Define a bijection σ on $\{0, 1, \dots, n-1\}$ such that, for $i \in I_0$, $\sigma(i) \in I_1$ and $(i, \sigma(i)]$ is a minimal span for G , and, for $j \in J_0$, $\sigma(j) \in J_1$ and $(\sigma(j), j]$ is a minimal span for H . Thus, $\{(i, \sigma(i)) : i = 0, 1, \dots, n-1\}$ is the collection of spans for X and $\{(\sigma(j), j) : j = 0, 1, \dots, n-1\}$ is the collection of spans for Y .

In the interpretation of [15, Section V, Figure 14], the set $\{(i, \sigma(i)) : i = 0, 1, \dots, n-1\}$ defines the positions of n nonattacking rooks on a chess board of size n , divided into $|I_1|$ black rooks above the main diagonal and $|J_1|$ white rooks below the main diagonal. Appendix C shows a rook configuration from a Suzuki function field.

A slightly different version of a subspace separating subset is a stopping set used in erasure decoding. Let $I \cup J$ be a partition of the coordinates and assume that the vector c is known in the positions I and erased in the positions J . A subset S of the dual code D is a stopping set for J if it is capable of recovering at least one of the erased symbols in J .

$$(\exists j \in J) D(I \cup j) \neq D(I) \Leftrightarrow (\exists j \in J) D(I \cup j) \cap S \neq D(I) \cap S. \quad (9)$$

4.2 Unit memory convolutional codes

The decomposition $X = X_0 + X_1$ has an interpretation as the *unwrapping* of the matrix X into two matrices X_0 and X_1 that define a unit-memory convolutional code via $y_t = u_t X_0 + u_{t-1} X_1$, with y_t and u_t vectors of length n indexed by $t \geq 0$.

$$\begin{bmatrix} X_0 & X_1 & & \\ & X_0 & X_1 & \\ & & \ddots & \ddots \end{bmatrix} \quad \begin{bmatrix} Y_0 & & \\ Y_1 & Y_0 & \\ & \ddots & \ddots \end{bmatrix}$$

For the codes below (X on the left and Y on the right) we extend the time t to all integers.

$$\begin{array}{cccc|cccccccc} 2 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 1 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 2 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & 2 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \quad \begin{array}{cccccccc|cccccccc} \cdot & 0 & 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 1 & 1 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 1 & 1 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

5 Duality for characteristic matrices

We first collect properties that hold for any pair of characteristic matrices X and Y . When we write $X = X_0 + X_1$ or $Y = Y_1 + Y_0$ it is implicitly understood that X_0 is upper triangular and X_1 strictly lower triangular, and Y_1 strictly upper triangular and Y_0 lower triangular.

Lemma 5.1. *For any pair of characteristic matrices X and Y , and for $X = X_0 + X_1$ and $Y = Y_1 + Y_0$, the row spaces of X and Y are orthogonal and the columns spaces of X_1 and Y_1 are orthogonal.*

$$XY^T = 0 \quad \text{and} \quad X_1^T Y_1 = 0. \quad (10)$$

The matrix X_1 has a full rank minor in the positions $J_0 \times J_1$. The matrix Y_1 has a full rank minor in the positions $I_0 \times I_1$. For reduced characteristic matrices and for semi-reduced characteristic matrices, the entries outside these minors are zero.

Proof. The row spaces of X and Y are those of the matrices G and H and $XY^T = 0$ follows from $GH^T = 0$. The matrix X_1 has zero rows in the positions I_0 and the matrix Y_1 has zero rows in the complementary positions J_0 . The nonzero rows of X_1 in the positions J_0 end in distinct positions of J_1 . The nonzero rows of Y_1 in the positions I_0 start in distinct positions of I_1 . The last claim follows from Definition 4.4 and Theorem 4.10. \square

The characteristic matrices X and Y in Theorem 4.10 satisfy the longer list of properties

$$X_0 Y_0^T = -I, \quad X_0 Y_1^T + X_1 Y_0^T = I, \quad X_1 Y_1^T = 0, \quad (11)$$

$$Y_0^T X_0 = -I, \quad Y_0^T X_1 + Y_1^T X_0 = I, \quad Y_1^T X_1 = 0, \quad (12)$$

and

$$Y^T X = 0. \quad (13)$$

We show that (11), (12) and (13) all follow from (10) and

$$X_0 Y_1^T + X_1 Y_0^T = I. \quad (14)$$

Lemma 5.2. *Let $X = X_0 + X_1$, with X_0 upper triangular and X_1 strictly lower triangular, and let $Y = Y_1 + Y_0$, with Y_1 strictly upper triangular and Y_0 lower triangular. If (10) and (14) hold then also (11), (12) and (13) hold. From (11) it follows that*

$$\left(\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & X_0 \end{array} \right) \left(\begin{array}{c|c} Y_0 & 0 \\ \hline Y_1 & Y_0 \end{array} \right)^T = \left(\begin{array}{c|c} -I & I \\ \hline 0 & -I \end{array} \right) \quad (15)$$

Proof. From $XY^T = 0$ and $X_0 Y_1^T + X_1 Y_0^T = I$, we obtain $X_0 Y_0^T + X_1 Y_1^T = -I$. Using that $X_0 Y_0^T$ is upper triangular and $X_1 Y_1^T$ is strictly lower triangular (11) and (15) follow. After writing (15) as

$$\left(\begin{array}{c|c} X_0 & X \\ \hline 0 & X_0 \end{array} \right) \left(\begin{array}{c|c} Y_0^T & Y_1^T \\ \hline 0 & Y_0^T \end{array} \right) = \left(\begin{array}{c|c} -I & 0 \\ \hline 0 & -I \end{array} \right).$$

the matrices on the left commute and (12) follows. Finally, after expanding $Y^T X$, $Y^T X = 0$ follows from (12). \square

The equality $XY^T = 0$ is clear. The row spaces of X and Y are those of the matrices G and H and are thus orthogonal. The equality $Y^T X = 0$ on the other hand, i.e. orthogonality of the column spaces of X and Y , is a property that does not hold for general choices of a characteristic matrix X for G and a characteristic matrix Y for H . It has the following trivial but important consequence for trellis constructions.

Lemma 5.3. *Assume that X and Y have maximal orthogonal column spaces. A selection of rows in X forms a basis for the row space of X if and only if the complementary rows in Y form a basis for the row space of Y .*

Proof. We avoid relying on matroid duality and write out the short proof: The selected rows in X form a basis if and only if there exists no nonzero column in X with zeros in the selected rows if and only if the complementary rows in Y form a basis. \square

The property $Y^T X = 0$, i.e. the orthogonality of column spaces, does not hold in general and is a consequence of $X_0 Y_1^T + X_1 Y_0^T = I$. The latter expresses a duality between rows of $(X_0|X_1)$ and rows of $(Y_1|Y_0)$, inner products are 1 for rows in the same position and 0 otherwise. The rows of $(X_0|X_1)$ form a basis for the coset scheme $(H|H)^\perp/(G|G)$ and those of $(Y_1|Y_0)$ a dual basis for the dual coset scheme $(G|G)^\perp/(H|H)$.

Definition 5.4. We say that characteristic matrices $X = X_0 + X_1$ and $Y = Y_1 + Y_0$ are in duality if $X_0 Y_1^T + X_1 Y_0^T = I$. In other words if, in addition to X and Y being characteristic matrices for the row spaces of orthogonal matrices G and H , the row spaces of $(X_0|X_1)$ and of $(Y_1|Y_0)$ are in duality as bases for the coset schemes $(H|H)^\perp/(G|G)$ and $(G|G)^\perp/(H|H)$.

The definition is justified by Equation (15) in Lemma 5.2, which shows that any characteristic matrix has a unique dual characteristic matrix, such that a left-ordered characteristic matrix is in duality with a right-ordered characteristic matrix. We still need to establish that this duality is the one that corresponds to trellis duality. This is done in the following theorem.

Theorem 5.5. *Trellises constructed via the Nori-Shankar span based BCJR construction from dual characteristic matrices X and Y are in duality.*

Proof. The first part of the claim is the rank duality: for a selection of rows in X that form a basis for the row space of X , the rows in the complementary positions in Y form a basis for the row space of Y . This was shown in Lemma 5.3 as an immediate consequence of $X^T Y = 0$. The second part is the duality of displacement matrices. We have to prove this duality for every pair of dual trellises obtained from X and Y . Displacement matrices are submatrices of $XY_1^T = X_0Y_1^T =: D$ and of $YX_1^T = Y_0X_1^T =: E$. For I a set of independent rows in X , the complement J is a set of independent rows in Y (by rank duality). The displacement matrices for the corresponding trellises are the submatrices $D|_{I \times J}$ and $E|_{J \times I}$. From $I \cap J = \emptyset$ and $D + E^T = X_0Y_1^T + X_1Y_0^T = I$ we see that

$$E[J \times I]^T = -D[I \times J] \quad (16)$$

which expresses duality of two tail-biting trellises as BCJR trellises. \square

The presence of the minus sign in (16) means that if we decide to use E and D to label vertices then syndromes will be accumulated in opposite directions for a BCJR trellis and its dual. A different approach to construct dual characteristic matrices and to prove trellis duality can be found in [11]. A core lemma in that paper is the following, stated here in a different form using the notation and concepts of this paper. We have also changed the roles of X and Y .

Lemma 5.6. ([11, Lemma 4.5]) *Given a characteristic matrix Y in right-ordered form, the top row c_0 of the dual characteristic matrix X (and thus the top row of the matrix X_0) is obtained as a solution to the system of equations*

$$Y_0 c_0^T = (1, 0, \dots, 0)^T.$$

Together with [11, Proposition 4.6], which states that every other row of X can be computed as the top row of a modified characteristic matrix determined by X , this yields a construction for the dual characteristic matrix. Clearly, in the given form the equation that is used to compute the top row of X matches $Y_0 X_0^T = -I$ (as before up to a change of sign in the scaling). The proof in [11] that the constructed dual characteristic matrix yields dual trellises is in several steps [11, Theorem 4.8, Proposition 4.9, Proposition 4.11].

Lemma 5.7. *The decomposition $I = X_0Y_1^T + X_1Y_0^T$ in (14) is a decomposition of orthogonal idempotents. For $D = X_0Y_1^T$ and $E = Y_0X_1^T$, $I = D + E^T$,*

$$D^2 = D, \quad E^2 = E, \quad DE^T = 0 \quad \text{and} \quad E^T D = 0.$$

It holds that $X = DX_0$ and $Y = EY_0$.

Proof. The orthogonality $DE^T = 0$ follows from $Y_1^T X_1 = 0$ in (10) and $E^T D = X_1(Y_0^T X_0)Y_1^T = 0$ from (11) and (12). For the relations involving D use

$$\begin{aligned} X_0 &= X_0 Y_1^T X_0 + X_1 Y_0^T X_0 = X_0 Y_1^T X_0 - X_1. \\ X &= X_0 + X_1 = X_0 Y_1^T X_0 = D X_0. \\ D &= X_0 Y_1^T = X Y_1^T = D X_0 Y_1^T = D^2. \end{aligned}$$

The relations involving E follow in the same way. □

5.1 On a conjecture by Koetter and Vardy

In the previous section we defined a duality for characteristic matrices and showed that it is compatible with both local trellis duality (where it agrees with the duality formulated in [10]) and with BCJR trellis duality (where it agrees with the displacement matrix duality from [17]). The dualities provide a bijection between characteristic matrices for G and those for its orthogonal matrix H . Koetter and Vardy had earlier proposed that there might be a duality between two special choices of characteristic matrices. In this section we adjust and then prove their proposed duality.

In [14] it was proposed to choose the lexicographically first one among generators with the same characteristic span. This defines a unique set of characteristic generators for a row space. It was then conjectured that the lexicographically first choices for a row space and its orthogonal dual would be in duality as characteristic matrices. Since then, examples have been found by Gleussen-Larssing and Weaver [10], [11] for which the claimed duality fails. Moreover, the authors provide a detailed duality that matches each characteristic matrix to a unique dual characteristic matrix. Both the construction of a dual characteristic matrix and the proof that it is dual to the original matrix are given in several steps and involve many technical details. In the previous section we used the unwrapped versions $(X_0|X_1)$ and $(Y_1|Y_0)$ to formulate an explicit duality combined with a shorter proof.

The original conjecture and subsequent work ignores the different direction of the time-axis in a trellis and its dual. After taking into account that lexicographically ordered has opposite meanings for a trellis and its dual trellis the conjecture is correct. The notion of lexicographical order that we use is any ordering of a field where the 0 element is minimal combined with an ordering of polynomials $\sum_k c_k x^k < \sum_k c'_k x^k$ whenever $c_k < c'_k$ for the largest k with $c_k \neq c'_k$. Row vectors in the expanded matrix $(X_0|X_1)$ follow the ordering (c_0, c_1, \dots) but row vectors in the dual expanded matrix $(Y_1|Y_0)$ follow the reverse ordering (\dots, d_1, d_0) . In both cases we order rows as polynomials. This leads to a selection of minimal characteristic generators that corresponds to that of a matrix in reduced minimal span form. We replace the original conjecture with the following theorem.

Theorem 5.8. *Reduced characteristic matrices are in duality.*

Proof. For a characteristic matrix X , the unwrapped form $(X_0|X_1)$ is in minimal span form with trailing pivots in the positions $I_1 \cup n + J_1$ and with full minor $A = (X_0|X_1)|_{I_1 \cup n + J_1}$ of the form $A = PL$. A reduced characteristic matrix X has X_1 with support in the positions $J_0 \times J_1$, where it has a minor of full rank (Lemma 5.1). The dual characteristic matrix Y , with unwrapped form $(Y_1|Y_0)$, satisfies $X_1 Y_1^T = Y_1^T X_1 = 0$ and therefore Y_1 has support in the positions $I_0 \times I_1$. The matrix $(Y_1|Y_0)$ is in minimal span form with leading pivots in the positions $I_1 \cup n + J_1$ and with full minor $B = (Y_1|Y_0)|_{I_1 \cup n + J_1}$ of the form $B = P'U$. The matrix $(X_0|X_1)$ is zero in the columns $n + I_1$ and the matrix $(Y_1|Y_0)$ is zero in the columns J_1 . Thus $(X_0|X_1)(Y_1|Y_0)^T = I$ becomes $AB^T = I$. So that $P' = P$ and $U = L^{-T}$. Moreover, since X is reduced, $A = L'P$. But then $AB^T = I$ shows that $B = L'^{-T}P = U'P$ and thus Y is reduced. \square

5.2 Column space characterization of duality

We give a characterization of duality that refers only to X and Y and not to their decompositions as sums of triangular matrices.

Lemma 5.9. *Different characteristic matrices X and X' for the same row space G have different column spaces.*

Proof. Let Y be the dual characteristic matrix for X , so that columns of X and Y are orthogonal. With Lemma 4.3, the matrices

$$\left(\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G \end{array} \right) \sim \left(\begin{array}{c|c} X'_0 & X'_1 \\ \hline 0 & G \end{array} \right)$$

are row equivalent, are both in left-ordered minimal span form, and share the same set of spans. One, X' , can be obtained from the other, X , via a set of upper triangular row operations. Only row operations are allowed that preserve the minimal span form. Thus a row operation that adds a word with span $(i, j]$ to another word is allowed only if the other word has span $(i', j']$ such that $(i, j] \subset (i', j']$. A change in column a affects only rows that contain a in their span. The corresponding rows in Y have spans that do not contain a . Thus these rows in Y are linearly independent by Lemma 4.12. And any difference between a column of X and a column of X' does not give a relation among rows of Y . This implies that any column of X' that is different from the corresponding column of X is not orthogonal to the column space of Y and therefore does not belong to the column space of X . \square

With the lemma, given a characteristic matrix Y for H , there exists a unique characteristic matrix X for G such that $Y^T X = 0$.

Theorem 5.10. *A pair of characteristic matrices X and Y , with maximal orthogonal row spaces, is in duality if and only if $Y^T X = 0$, that is if and only if X and Y have orthogonal column spaces.*

Proof. Duality is defined by $X_0 Y_1^T + X_1 Y_0^T = I$ and this implies that $Y^T X = 0$. This proves the only if part. The if part follows from the lemma. \square

5.3 Displacement matrices

We give further properties of the matrices $D = X_0 Y_1^T$ and $E = Y_0 X_1^T$, that have a role as displacement matrix in the BCJR constuction of [17]: Let the rows I of X be chosen as edge labelings for generators of a label code, and let the complementary rows J of Y be chosen as edge labelings for generators of a dual label code. For row i of X , $i \in I$, the corresponding row i of the submatrix $E[J \times I]^T$ is selected as the initial vertex label of the path. Similarly, for row j of Y , $j \in J$, the corresponding row j of the submatrix $D[I \times J]^T$ is selected as the initial vertex label of the path. The labeling of the remaining vertices in the label code proceeds as in the BCJR construction via the accumulation of partial syndromes.

For characteristic matrices X and Y , $XY^T = 0$ and $X_1^T Y_1 = 0$. Moreover, for matrices in duality, $X^T Y = 0$ and $X_1 Y_1^T = 0$. In general the minors $J_0 \times J_1$ of X_1 and $I_0 \times I_1$ of Y_1 are invertible. The matrix X_1 is zero in rows I_0 , and the matrix Y_1 is zero in rows J_0 . We write this as

$$D_0 X_1 = 0, \quad E_0 Y_1 = 0,$$

for diagonal 0,1-matrices D_0 and E_0 , with $D_0 = 1$ in the positions I_0 and $E_0 = 1$ in the positions J_0 . So that $D_0 + E_0 = I$.

Lemma 5.11. *The nonzero columns I_0 of D form a basis in systematic form for the column space of X and the nonzero columns J_0 of E form a basis in systematic form for the column space of Y . Equivalently,*

$$DD_0 = D, \quad D_0 D = D_0, \quad D^T Y = 0 \tag{17}$$

$$EE_0 = E, \quad E_0 E = E_0, \quad E^T X = 0 \tag{18}$$

Proof.

$$DD_0 = D(I - E_0) = D - X_0 Y_1^T E_0 = D,$$

$$D_0 D = D_0(I - X_1 Y_0^T) = D_0,$$

$$(D^T)Y = Y_1 X^T Y = 0.$$

Similar for (18). □

From (17) and (18) we see that

$$X = (D + E^T)X = DX = D(D_0 X). \tag{19}$$

Here $D_0 X$ has as nonzero rows the conventional generators of X . And another interpretation for D is that it describes the rows of X as linear combinations of its conventional generators.

Let S be the permutation matrix that shifts row vectors to the right and column vectors upwards: $e_0 S = e_1, \dots, e_{n-1} S = e_0$, and $S e_1^T = e_0^T, \dots, S e_0^T = e_{n-1}^T$. Let $X^S = S X S^T$ be the conjugate of a left-ordered characteristic matrix X for G . Then X^S is a left-ordered characteristic matrix for $G S^T$. If Y is a right-ordered characteristic matrix for H then Y^S is a right-ordered characteristic matrix for $H S^T$. For $D(X) = X_0 Y_1^T = X X_0^{-1}$ it is in general not true that $D(X^S) = D(X)^S$.

Lemma 5.12. *Let X have left column x and Y left column y . Then*

$$D(X^S)^{S^T} = D(X) + xy^T, \quad E(Y^S)^{S^T} = E(Y) - yx^T.$$

Proof. Let $Y = \left(\begin{array}{c|c} d & d_0 \\ \hline d_1 & B \end{array} \right)$. Then

$$D(X) = XY_1^T = X \left(\begin{array}{c|c} 0 & d_0 \\ \hline 0 & B_1 \end{array} \right)^T, \quad Y^S = \left(\begin{array}{c|c} B & d_1 \\ \hline d_0 & d \end{array} \right),$$

And

$$\begin{aligned} D(X^S)^{S^T} &= (X^S)^{S^T} ((Y^S)_1^T)^{S^T} = XS^T \left(\begin{array}{c|c} B_1 & d_1 \\ \hline 0 & 0 \end{array} \right)^T S = X \left(\begin{array}{c|c} 0 & 0 \\ \hline d_1 & B_1 \end{array} \right)^T \\ &= D(X) + X \left(\begin{array}{c|c} d & 0 \\ \hline d_1 & 0 \end{array} \right)^T - X \left(\begin{array}{c|c} d & d_0 \\ \hline 0 & 0 \end{array} \right)^T = D(X) + xy^T. \end{aligned}$$

The claim for $E(Y^S)$ follows with $D + E^T = I$. \square

Theorem 5.13. *Let $S(Y^T|X)$ be the label code representing the rows of X as a BCJR trellis with vertex labelings from the column space of Y . Let N_0, N_1, \dots, N_n be the submatrices of $S(Y^T|X)$ representing the vertex labelings. The characterization of $N_0 = E^T$ by Lemma 5.11 as a matrix with nonzero rows in systematic form orthogonal to columns of X applies to all N_0, N_1, \dots, N_n .*

Example 5.14.

$$\begin{aligned} Y^T &= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix} & X &= \begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 2 \end{bmatrix} \\ X^T &= \begin{bmatrix} 2 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 2 \end{bmatrix} & Y &= \begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \\ S(Y^T|X) &= \left[\begin{array}{cccc|c|cccc|c|cccc|c|cccc} 0 & 0 & 0 & 0 & 2 & 1 & 1 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 1 & 2 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 1 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 2 & 0 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array} \right] \\ S(X^T|Y) &= \left[\begin{array}{cccc|c|cccc|c|cccc|c|cccc} 1 & 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 2 & 1 \\ 0 & 1 & 2 & 2 & 1 & 2 & 1 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 1 & 2 & 1 & 1 & 1 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

5.4 Decompositions of the characteristic matrix

In general, the matrices X and Y_1 have minors of maximal full rank in the positions $I_0 \times I_1$. And the matrices Y and X_1 have minors of maximal full rank in the positions $J_0 \times J_1$. From Lemma 5.7, $X = X_0 Y_1^T X_0$ and $Y = Y_0 X_1^T Y_0$. With X_0 and Y_0 triangular and invertible, we recognize the format of a Bruhat decomposition (Section 2.1): Submatrices of X and Y_1^T of the same size and positioned in the south west corner are of same rank, and submatrices of Y and X_1^T of the same size and positioned in the north east corner are of same rank.

For reduced characteristic matrices and for the characteristic matrices of Theorem 4.10, X_1 and Y_1 are zero outside the invertible minors. This is not the case in general. In general Y_1 can be nonzero outside the columns I_1 and the matrix X_1 can be nonzero outside the columns J_1 .

Example 5.15.

$$G = H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ \cdot & 1 & 1 & 1 \\ \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & 1 \end{bmatrix} + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot \\ 1 & 1 & 1 & \cdot \end{bmatrix} \quad \begin{array}{l} J_0 = \{2, 4\} \\ J_1 = \{1, 3\} \end{array}$$

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \cdot & 1 & 1 & 1 \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} + \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot \\ 1 & 1 & 1 & \cdot \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \begin{array}{l} I_0 = \{1, 3\} \\ I_1 = \{2, 4\} \end{array}$$

In general

$$X = X[n \times I_1] \times X[I_0 \times I_1]^{-1} \times X[I_0 \times n],$$

$$Y = Y[n \times J_1] \times Y[J_0 \times J_1]^{-1} \times Y[J_0 \times n].$$

Let $W^T = X[n \times I_1]$, $G = X[I_0 \times n]$, $V^T = Y[n \times J_1]$, and $H = Y[J_0 \times n]$. In fact, as submatrices of X and Y containing the conventional generators, $G = G_{01}$ and $H = H_{10}$ are in reduced minimal span form. The effect of the invertible minor is to bring W^T and V^T in systematic form in the positions I_0 and J_0 . Thus $X = W_0^T G_{01}$ and $Y = V_1^T H_{10}$.

In the special case where X_1 and Y_1 are zero outside their full rank minors, we have

$$Y_1 = D_0 Y_1 D_1, \quad X_1 = E_0 X_1 E_1,$$

with D_1 and E_1 diagonal 0, 1,-matrices with support in I_1 and J_1 , respectively. This is the case for reduced characteristic matrices and for the characteristic matrices in Theorem 4.10. From $I = X_0 Y_1^T + X_1 Y_0^T = X Y_1^T + X_1 Y^T$, $D_0^2 = (D_0 X D_1)(D_0 Y_1 D_1)^T$ and thus

$$X[I_0 \times I_1] \times Y_1[I_0 \times I_1]^T = I.$$

5.5 The transpose of a characteristic matrix

By Proposition 4.11 the row span of a characteristic matrix X is $|J_1|n$.

Lemma 5.16. *Let X be a left-ordered characteristic matrix with spanlength $|J_1|n$. The rows in X^T have spans ending on the diagonal. The spanlength of X^T is at least $|J_1|n$, with equality if and only if X is reduced. Let Y be a right-ordered characteristic matrix with spanlength $|I_1|n$. The rows in Y^T have spans starting on the diagonal. The spanlength of Y^T is at least $|I_1|n$, with equality if and only if Y is reduced.*

Proof. Let $(i_0, i_1]$ be the span of a row i_0 in X . Then the span of row i_1 in X^T includes $(i_0, i_1]$. The span is equal to $(i_0, i_1]$ for all i_0 if and only if X is reduced. Similar for Y . \square

Theorem 5.17. *The transpose of a characteristic matrix is again a characteristic matrix if and only if the characteristic matrix is reduced.*

(X) *The transpose of a reduced left-ordered characteristic matrix X is a right-ordered characteristic matrix X^T . A minimal span form for the row space of X is given by $X[I_0 \times n]$ and a minimal span form for the row space of X^T by $X[n \times I_1]^T$.*

(Y) *The transpose of a reduced right-ordered characteristic matrix Y is a left-ordered characteristic matrix Y^T . A minimal span form for the row space of Y is given by $Y[J_0 \times n]$ and a minimal span form for the row space of Y^T by $Y[n \times J_1]^T$.*

Proof. The lemma shows that the spanlength of a characteristic matrix increases when passing to the transpose unless the matrix is reduced. This proves the only if part. When the matrix is reduced taking the transpose preserves the set of characteristic spans and the spanlength, and the transpose is again a characteristic matrix. \square

To a reduced pair of dual characteristic matrices X and Y (with row spaces $G \perp H$) thus corresponds a reduced pair of dual characteristic matrices Y^T and X^T (with row spaces $V \perp W$). The matrices G and W are of the same size and share the same characteristic spans. The same for the matrices H and V .

Example 5.18. There are three binary row spaces with characteristic spans $(0, 3]$, $(1, 0]$, $(2, 4]$, $(3, 1]$, $(4, 2]$. They are

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad G' = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \quad G'' = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

By the above $\{X^T(G), X^T(G'), X^T(G'')\} = \{Y(G), Y(G'), Y(G'')\}$. In fact, $X^T(G) = Y(G)$, $X^T(G') = Y(G'')$, $X^T(G'') = Y(G')$. Thus the matrices G' and G'' describe the column spaces of each others reduced characteristic matrices.

Let X have row space generated by G and column space generated by W^T . The construction of a KV-trellis from a characteristic matrix requires a choice of independent generators from the rows of X . A set of rows in X is independent if and only if the corresponding columns in W are independent. Thus the set of possible choices is in bijection with the set of independent full minors of W , that is with the bases for the matroid of W .

Example 5.19. The only matrix in the previous example that admits a minimal trellis with $|V_i| \leq 2$ for all $i \in \{0, 1, 2, 3, 4\}$ is G' . Namely such a trellis can only be formed with generators having spans $(2, 4]$ and $(4, 2]$. Those generators are independent if and only if, for $X^T(G) = Y(W)$, the columns 2 and 4 in W are linearly independent. This occurs only if $W = G''$. The required trellis is generated by the rows 00111 and 11101.

For reduced characteristic matrices X and Y , the row spaces $G = X[I_0 \times n]$, $W = X[n \times I_1]^T$, $H = Y[J_0 \times n]$, and $V = Y[n \times J_1]^T$ are all in minimal span form, with $G = G_{01}$, $W = W_{10}$, $H = H_{10}$, and $V = V_{01}$. The following four matrices are in reduced minimal span form

$$\left[\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G \end{array} \right] \quad \left[\begin{array}{c|c} W & 0 \\ \hline X_1^T & X_0^T \end{array} \right] \quad \left[\begin{array}{c|c} H & 0 \\ \hline Y_1 & Y_0 \end{array} \right] \quad \left[\begin{array}{c|c} Y_0^T & Y_1^T \\ \hline 0 & V \end{array} \right]$$

Moreover, for the systematic forms G_1, W_0, H_0, V_1 ,

$$X = W_0^T G_{01} \quad X^T = G_1^T W_{10} \quad Y = V_1^T H_{10} \quad Y^T = H_0^T V_{01}$$

Circulant matrices are special cases of reduced characteristic matrices. In this case there is a clear relation between the row space and the column space of the circulant matrix.

Proposition 5.20. *Let C' denote the reverse ordered version of a cyclic code C . Then $X(C)^T = Y(C')$ and $Y(C)^T = X(C')$.*

References

- [1] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Information Theory*, IT-20:284–287, 1974.
- [2] A. R. Calderbank, G. David Forney, Jr., and Alexander Vardy. Minimal tail-biting trellises: the Golay code and more. *IEEE Trans. Inform. Theory*, 45(5):1435–1455, 1999.
- [3] David Conti and Nigel Boston. On the algebraic structure of linear tail-biting trellises. *IEEE Trans. Inform. Theory*, 61(5):2283–2299, 2015.
- [4] Iwan Duursma, Radoslav Kirov, and Seungkook Park. Distance bounds for algebraic geometric codes. *J. Pure Appl. Algebra*, 215(8):1863–1878, 2011.
- [5] Iwan M. Duursma. Two-point coordinate rings for GK-curves. *IEEE Trans. Inform. Theory*, 57(2):593–600, 2011.
- [6] Iwan M. Duursma and Seungkook Park. Delta sets for divisors supported in two points. *Finite Fields Appl.*, 18(5):865–885, 2012.
- [7] A. Yu. Erëmin and L. Yu. Kolotilina. The Bruhat decomposition and the solution of systems of linear algebraic equations. *Sov. J. Numer. Anal. Math. Modelling*, 2(6):421–436, 1987.

- [8] G. David Forney, Jr. Codes on graphs: normal realizations. *IEEE Trans. Inform. Theory*, 47(2):520–548, 2001.
- [9] G. David Forney, Jr. Minimal realizations of linear systems: the “shortest basis” approach. *IEEE Trans. Inform. Theory*, 57(2):726–737, 2011.
- [10] Heide Gluesing-Luerssen and Elizabeth A. Weaver. Linear tail-biting trellises: characteristic generators and the BCJR-construction. *IEEE Trans. Inform. Theory*, 57(2):738–751, 2011.
- [11] Heide Gluesing-Luerssen and Elizabeth A. Weaver. Characteristic generators and dualization for tail-biting trellises. *IEEE Trans. Inform. Theory*, 57(11):7418–7430, 2011.
- [12] G. David Forney Jr. Coset codes-ii: Binary lattices and related codes. *IEEE Transactions on Information Theory*, 34(5):1152–1187, 1988.
- [13] Ralf Koetter and Alexander Vardy. On the theory of linear trellises. In *Information, coding and mathematics*, volume 687 of *Kluwer Internat. Ser. Engrg. Comput. Sci.*, pages 323–354. Kluwer Acad. Publ., Boston, MA, 2002.
- [14] Ralf Koetter and Alexander Vardy. The structure of tail-biting trellises: minimality and basic principles. *IEEE Trans. Inform. Theory*, 49(9):2081–2105, 2003.
- [15] Frank R. Kschischang and Vladislav Sorokine. On the trellis structure of block codes. *IEEE Trans. Inform. Theory*, 41(6, part 2):1924–1937, 1995.
- [16] Robert J. McEliece. On the BCJR trellis for linear block codes. *IEEE Trans. Inform. Theory*, 42(4):1072–1092, 1996.
- [17] Aditya V. Nori and Priti Shankar. Unifying views of tail-biting trellis constructions for linear block codes. *IEEE Trans. Inform. Theory*, 52(10):4431–4443, 2006.
- [18] Aditya Vithal Nori. *Unifying views of tail-biting trellises for linear block codes*. PhD thesis, Indian Institute of Science Bangalore, 2005.
- [19] G. Solomon and H. C. A. van Tilborg. A connection between block and convolutional codes. *SIAM J. Appl. Math.*, 37(2):358–369, 1979.
- [20] T. A. Springer. *Linear algebraic groups*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, second edition, 2009.
- [21] Alexander Vardy. Trellis structure of codes. In *Handbook of coding theory, Vol. I, II*, pages 1989–2117. North-Holland, Amsterdam, 1998.
- [22] Elizabeth A Weaver. *Minimality and duality of tail-biting trellises for linear codes*. PhD thesis, University of Kentucky, 2012.
- [23] Jack K. Wolf. Efficient maximum likelihood decoding of linear block codes using a trellis. *IEEE Trans. Information Theory*, IT-24(1):76–80, 1978.

A A dual pair of reduced characteristic matrices

Reduced characteristic matrices X and Y for a pair of binary orthogonal matrices G and H .

$$\begin{aligned}
 \left[\begin{array}{c|c} X_0 & X_1 \\ \hline 0 & G \end{array} \right] &= \left[\begin{array}{cc|cc|cccc} 1 & 1 & 1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 1 & 0 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 1 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 & 0 & 0 & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 \end{array} \right] \\
 &\perp \left[\begin{array}{ccc|cc} 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & \cdot \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{array} \right] = [H | H]. \\
 [G | G] &= \left[\begin{array}{cc|ccc|ccc} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & \cdot & \cdot \\ \cdot & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ \cdot & \cdot & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{array} \right] \\
 &\perp \left[\begin{array}{cc|ccc|ccc} 0 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 0 & 1 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 0 & 1 & 1 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 0 & 0 & 1 & 1 & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & 0 & 0 & 1 & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 0 & 1 & 1 & 1 \end{array} \right] = \left[\begin{array}{c|c} H & 0 \\ \hline Y_1 & Y_0 \end{array} \right] \\
 X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ \cdot & 1 & 1 & 0 & 1 \\ \cdot & \cdot & 1 & 1 & 0 \\ \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot \\ 1 & 0 & 0 & 0 & \cdot \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} (0,2) \\ (1,4) \\ (2,3) \\ (3,1) \\ (4,0) \end{matrix} \\
 Y = \begin{bmatrix} \cdot & 0 & 1 & 1 & 1 \\ \cdot & \cdot & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} + \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot \\ 0 & 1 & 1 & \cdot & \cdot \\ 0 & 1 & 1 & 1 & \cdot \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{matrix} (2,0) \\ (4,1) \\ (3,2) \\ (1,3) \\ (0,4) \end{matrix}
 \end{aligned}$$

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \perp H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

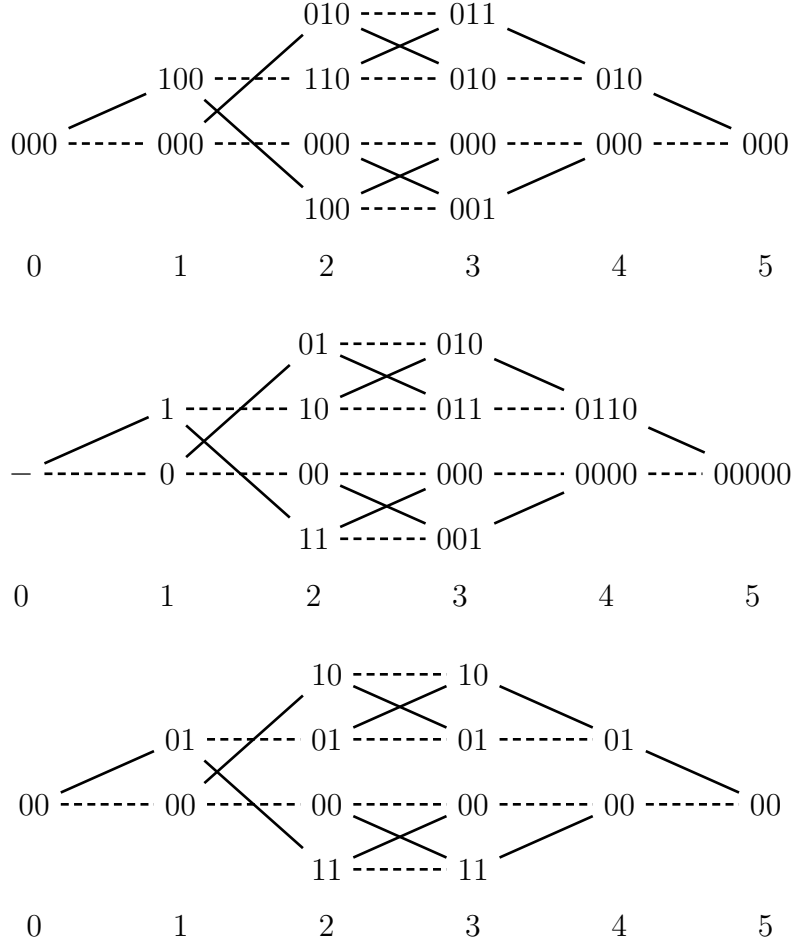


Figure 3: Three different labelings for the minimal conventional trellises of the row space of G : vertex labeling by information symbols (top), by codeword symbols (middle), and by syndromes (bottom)

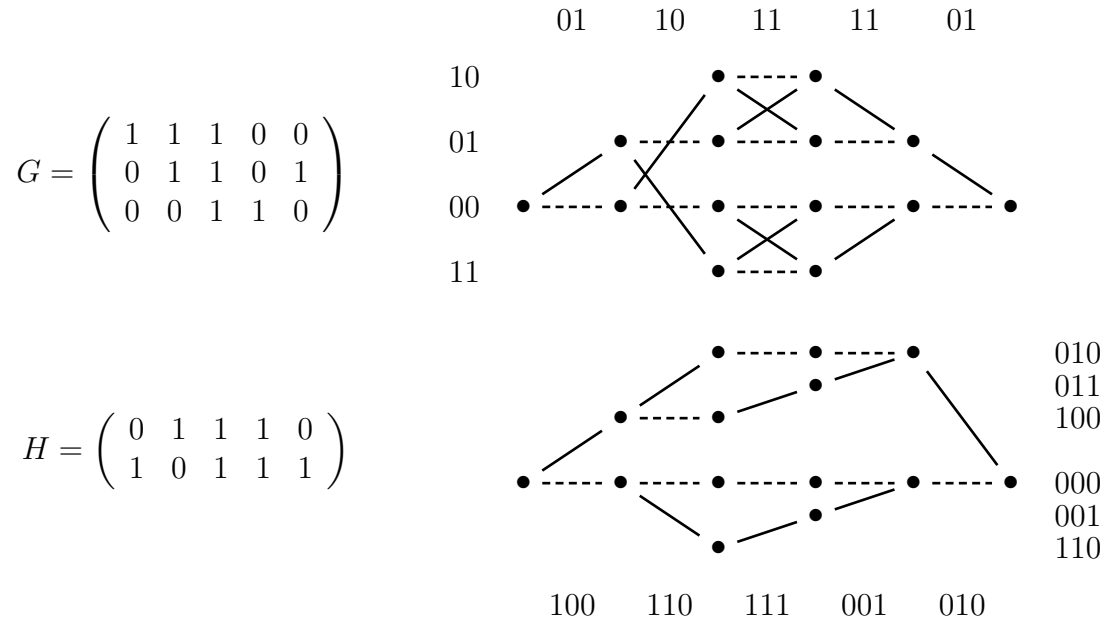


Figure 4: A pair of dual conventional trellises, based on rows $X_{1,2,3}$ and $Y_{4,5}$

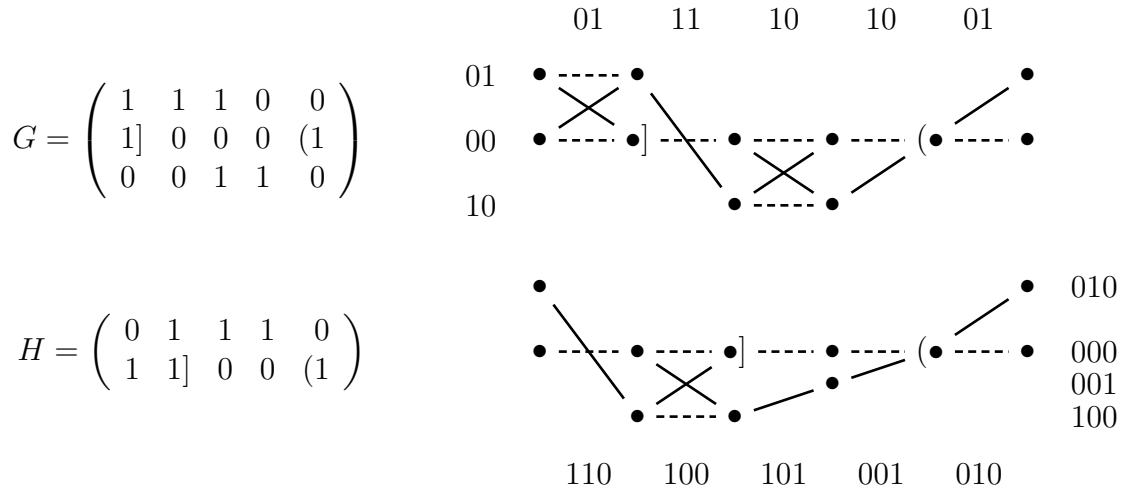


Figure 5: A pair of dual tailbiting trellises, based on rows $X_{1,5,3}$ and $Y_{4,2}$

Label codes for KV-trellises constructed from X and Y using the span based BCJR construction.

$$Y^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$X^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad Y = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$$S(Y^T|X) = \left[\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Label codes for the dual trellises defined with rows 1, 3, 5 of X and rows 2, 4 of Y .

[illegible]

B A reduced characteristic matrix for the Golay code

The binary Golay code of length 24 is generated by the following twelve row vectors over $GF(4) = \{0, 1, a, b\}$ after the concatenation $0 \mapsto 00, 1 \mapsto 11, a \mapsto 01, b \mapsto 10$.

$$\begin{bmatrix} 1 & a & b & 1 & b & a & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a & b & 1 & a & 1 & b & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & b & 1 & a & 1 & b & a & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & a & b & 1 & b & a & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & a & b & 1 & b & a & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a & b & 1 & a & 1 & b & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & b & 1 & a & 1 & b & a \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a & b & 1 & b & a & 1 \\ b & a & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & a & b & 1 \\ 1 & b & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a & b & 1 & a \\ a & 1 & b & a & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & b & 1 \\ 1 & b & a & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a & b \end{bmatrix}$$

Every even row is a scalar multiple of the preceding odd row and every odd row is obtained with an obvious symmetry from the preceding even row. If we apply the concatenation in combination with a transposition of symbols $4k+2$ and $4k+4$, i.e.

$$\begin{aligned} ab &\mapsto 0110 \mapsto 0011, & b1 &\mapsto 1011 \mapsto 1110, & 1a &\mapsto 1101 \mapsto 1101, \\ ba &\mapsto 1001 \mapsto 1100, & 1b &\mapsto 1110 \mapsto 1011, & a1 &\mapsto 0111 \mapsto 0111, \end{aligned}$$

then the result is a generator matrix G for the Golay code with optimal minimal spanlength $108 = 12 \cdot 9$. The generator matrix is of the form (1) in [2] and agrees with the generator matrix in (4) of that paper after a permutation $(1\ 2)(7\ 8)(9\ 10)(15\ 16)(17\ 18)(23\ 24)$. The reduced characteristic matrix X for G is of the form

$$X = \begin{bmatrix} A & B & C \\ C & A & B \\ B & C & A \end{bmatrix}$$

with $(A|B|C)$ equal to

$$\left[\begin{array}{cccccccc} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \middle| \begin{array}{cccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{array} \middle| \begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{array} \right]$$

The odd rows have spanlength 9 and the even rows have spanlength 15.

C Nonattacking rooks for function fields

Minimal spans for characteristic generators have an analogue in discrepancies for function fields.

Given a field K of algebraic functions and two rational points (places), a function in K has span (i, j) if it has a pole of order j at the first point, no other poles, and a zero of order i at the second point. For each $j \in \mathbb{Z}$, let $\sigma(j) \in \mathbb{Z}$ be maximal such that there exists a function with span (i, j) . This defines a pair $(\sigma(j), j)$. The pairs fill the plane with an infinite set of nonattacking rooks. let $n > 0$ be minimal such that there exists a function with $(i, j) = (n, n)$. Then the pattern is periodic with period n .

For the function field $F(x, y)$, defined with equation $y^8 + y = x^{10} + x^3$ over the field of eight elements, $n = 13$. There exist functions with minimal spans

$$(i, j) \in \{(0, 0), (1, 8), (2, 16), (3, 10), (4, 18), (5, 12), (6, 20), \\ (7, 28), (8, 22), (9, 30), (10, 24), (11, 32), (12, 40)\}.$$

Figure 6 pictures the minimal spans as nonattacking rooks on a chess board of size 13 with labeling $\{0, 1, \dots, 11, 12\} \times \{1, 2, \dots, 12, 0\}$ modulo 13.

.	0
.	1
.	.	2
.	1
.	.	.	.	2
.	1	.	.
.	2
.	3
.	2
.	.	.	3
.	2	.	.	.
.	3
4

Figure 6: Nonattacking rooks for the Suzuki curve over the field of eight elements.